IN THE CLAIMS

1-32. (Canceled)

33. (New) A method for satisfying a Quality of Service (QoS) contract with an initiator, comprising:

receiving a request from an initiator in a first time less than or equal to an ordinal number times an arrival interval to satisfy an arrival model, wherein the ordinal number signifies a position of the request among a group of requests;

returning the request that has been serviced to the initiator in a second time less than or equal to a constant term plus the ordinal number times a service interval to satisfy a service model; and

measuring the first and second time at a boundary between the initiator and an interconnect.

34. (New) The method of claim 33, further comprising:

determining whether the arrival model and the service model have been satisfied in order to satisfy the QoS contract.

35. (New) The method of claim 33, further comprising:

satisfying the QoS contract using the initiator, a target, the interconnect, and a QoS unit that are all located on a chip.

36. (New) The method of claim 35, wherein the service interval is greater than the arrival interval.

37. (New) The method of claim 33, wherein the service interval is equal to the arrival interval.

38. (New) The method of claim 33, further comprising:

determining the constant term specifically for the group of requests.

39. (New) The method of claim 38, wherein the arrival interval is a predetermined accepted interval for requested arrivals.

40. (New) A system, comprising:

an interconnect coupled between an initiator and a target, wherein an arrival model is satisfied by the initiator if a request arrives before a first time less than or equal to an ordinal number times an arrival interval, wherein the ordinal number signifies a position of the request among a group of requests;

a Quality of Service (QoS) logic unit coupled to the interconnect and is configured to satisfy a service model if the initiator satisfies the arrival model, wherein the logic unit satisfies the service model if the request is serviced before a second time less than or equal to a constant term plus the ordinal number times a service interval; and

wherein the Quality of Service (QoS) logic unit is configured to satisfy a QoS contract and is located at a boundary between the initiator and the interconnect and the QoS contract is based on the arrival model and the service model being satisfied.

41. (New) The system of claim 40, wherein the service interval is greater than the arrival interval.

42. (New) The system of claim 40, wherein the constant term is determined specifically for the group of requests.

43. (New) The system of claim 42, wherein the QoS unit is coupled between the target and the interconnect.

44. (New) The method of claim 40, wherein the QoS contract is satisfied using the initiator, the target, the interconnect, and the logic that are all located on a chip.

45. (New) The system of claim 44, wherein the QoS unit is part of the interconnect.

46. (New) The system of claim 40, wherein the QoS unit is part of the target.

47. (New) A method for tracking service for two or more threads, comprising:

maintaining a first allocation count for a first thread to track whether the first thread is being serviced by a target;

establishing an adjustable positive limit having a first value for the first allocation count that establishes an initial maximum amount the first allocation count may go up to; and

raising the adjustable positive limit for the first allocation count to a second value when the first value for the first allocation count has been reached and the first thread has yet to have a request fulfilled from the target during a regular interval of time.

48. (New) The method of claim 47, further comprising:

reducing the positive limit from the second value to a third value for the first allocation count at a time when a second allocation count receives a credit and is positive.

49. (New) The method of claim 48, wherein the positive limit is not reduced below the first value.

50. (New) The method of claim 47, further comprising:

crediting the first and second allocation count at a regular interval;

debiting the first allocation count when the first thread is serviced; and debiting the second allocation count when a second thread is serviced; and establishing a negative limit having a fourth value for the first allocation count.

51. (New) The method of claim 47, wherein raising comprises increasing the positive limit by an amount proportional to an allocation rate of the first thread.

52. (New) The method of claim 51, wherein the amount proportional to the allocation rate of the first thread comprises a number of times the second thread was serviced by

the target during the regular interval of time multiplied by the allocation rate of the first thread.

53. (New) The method of claim 50, further comprising:

determining whether the first thread will be serviced according to the first allocation count.

54. (New) The method of claim 53, further comprising:

servicing the first thread instead of the second thread when the second allocation count is less than the first allocation count.

55. (New) The method of claim 47, wherein the second thread is a high-priority thread.

56. (New) The method of claim 55, wherein the first thread is a bandwidth-allocation thread.

57. (New) A system, comprising:

an interconnect coupled to an initiator and a target;

a first logic coupled to the interconnect, and configured to maintain a first allocation count for a first thread from the initiator to track bandwidth usage by the first thread;

a second logic coupled to the first logic, and configured to establish an adjustable positive limit having a first value for the first allocation count that establishes an initial maximum amount the first allocation count may go up to; and

a third logic coupled to the second logic, and configured to raise the adjustable positive limit from the first value to a second value when the first value for the first allocation count has been reached, then the first thread has not been service during a predetermined interval of time, and the first thread has one or more requests waiting to be serviced from the target.

58. (New) The system of claim 57, further comprising:

a fourth logic coupled to the second logic, and configured to reduce the raised adjustable positive limit to a third value when a second allocation count receives a credit and is positive.

59. (New) The system of claim 57, wherein the positive limit is raised by an amount proportional to an allocation rate of the first thread.

60. (New) The system of claim 57, wherein the first thread is a bandwidth allocation thread.

61. (New) The system of claim 60, further comprising:

a fifth logic coupled to the interconnect and the first logic, the fifth logic is configured to satisfy a service model if an arrival model is satisfied by the initiator, and wherein the fifth logic uses the first allocation count to satisfy the service model.

62. (New) The system of claim 61, wherein the arrival model is satisfied by the initiator if a request arrives before a first time less than or equal to an ordinal number times an arrival interval, wherein the ordinal number signifies a position of the request among a group of requests; and

wherein the service model is satisfied if the request is serviced before a second time less than or equal to a constant term plus the ordinal number times a service interval.